AWARD NUMBER: W81XWH-14-1-0152

TITLE: Scaffold Attachment Factor B1: A Novel Chromatin Regulator of Prostate Cancer Metabolism

PRINCIPAL INVESTIGATOR: SUNGYONG YOU PhD

CONTRACTING ORGANIZATION: Cedars-Sinai Medical Center
Frederick, MD 21702

REPORT DATE: August 2015

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
August 2015	Annual	1 Aug 2014 - 31 Jul 2015

4. TITLE AND SUBTITLE

Scaffold Attachment Factor B1: A Novel Chromatin Regulator of Prostate Cancer Metabolism

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-14-1-0152

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Sungyong You, Jayoung Kim, Michael R Freeman.

E-Mail:Sungyong.You@cshs.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

CEDARS-SINAI MEDICAL CENTER
8700 BEVERLY BLVD
LOS ANGELES CA 90048-1804

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Progress was made in four major areas of study in the first year of the funding period. First, I determined the overlapping target genes of SAFB1 and AR. Second, I identified a SAFB1-dependent gene signature in clinical specimens using a large prostate cancer transcriptome dataset I assembled. Interestingly, the enriched cellular processes were associated with aggressive cancer behavior, including steroid hormone response, blood vessel formation, and RNA processing. Third, I found I could classify prostate cancer subtypes into three distinct categories, using only transcriptome data. One of these subtypes is functionally not possible to characterize, based on currently published profiling and genomic analyses. Fourth, I searched for possible regulators of this newly identified subtype using an integrative network analysis of hundreds of castration-resistant PC RNA expression profiles. This analysis revealed a set of key transcription factors (TFs) that appear to play a major role in metastatic castration-resistant prostate cancer (mCRPC). Among these TFs, we identified ONECUT2, which appears to be a master regulator of the novel prostate cancer subtype in our classification scheme.

15. SUBJECT TERMS
Systems Biology, SAFB1, Prostate Cancer, Transcriptome

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	19	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. INTRODUCTION

Prostate cancer (PC) is a leading cause of death from cancer and there is no treatment for castration-resistant metastatic disease (CRPC) that substantially prolongs life. Recent studies on humans and laboratory models have provided evidence that high circulating cholesterol is a risk factor for aggressive PC[1-5]. We recently discovered that a protein, scaffold attachment factor B1 (SAFB1), is a novel regulator of the androgen receptor and other proteins associated with prostate cancer progression to end-stage disease[6]. The purpose of my research in this project is to identify and functionally characterize the gene regulatory networks controlled by SAFB1 in human PC cells.

This project is testing the hypotheses that ***(1) SAFB1 regulates a transcriptional program that leads to PC progression when perturbed by SAFB1 loss; and that (2) down-regulation of SAFB1 promotes CRPC in part through upregulation of cholesterol-dependent intracrine androgen signaling.*** To this end we performed chromatin immunoprecipitation-next generation DNA sequencing (ChIP-seq) and integrative network modeling to identify the SAFB1 cistrome and the extent of transcriptional collaboration of SAFB1, AR, and EZH2 in PC cells. These studies have been aided through the use of a large integrated transcriptome database of prostate cancer gene expression profiles of human tumors that I recently assembled, called the prostate cancer transcriptome atlas (PCTA). During the second year of the funding period we will test whether cholesterol alters intracrine androgen mechanisms in a SAFB1-dependent manner.

Our specific aims have *not* modified from those stated in the original application.

Specific Aim 1. To characterize the SAFB1 cistrome in prostate cancer cells and to determine the metabolic and biologic effects of SAFB1 loss.
Specific Aim 2. To test whether cholesterol alters intracrine androgen mechanisms in a SAFB1-dependent manner.

2. KEYWORDS

Systems Biology, SAFB1, Prostate Cancer, Transcriptome

3. ACCOMPLISHMENTS

What were the major goals of the project?

Training Goal 1: Training and educational development in prostate cancer research
Milestone: Presentation of project data at a national meeting
Target months: 24
Percentage of completion: 100%

Research Goal 1: To characterize the SAFB1 cistrome in prostate cancer cells and to determine the metabolic and biologic effects of SAFB1 loss.
Milestones:
1) Characterization of SAFB1 cistrome in the presence- or absence of dihydrotestosterone (DHT).
2) Determination of the overlapping target genes or sub-network between SAFB1 and AR or EZH2 and the genes or pathways involved in sterol metabolism and chromatin regulation.
3) Determination of the genes or pathways strongly associated with SAFB1 regulation and PC progression.
Target months: 12
Percentage of completion: 100%

Research Goal 2: To test whether cholesterol alters intracrine androgen mechanisms in a SAFB1-dependent manner.
Milestones:
1) Identification of critical regulatory nodes in the androgen metabolism network.

2) Characterization of the involvement of SAFB1 regulation of the UGT2B gene family, androgen metabolism, and downstream effects relevant to disease progression.
Target months: 24
Percentage of completion: 20%

What was accomplished under these goals?

These studies have identified novel links between the AR and EZH2 oncoproteins and genes that regulate sterol metabolism in CRPC. Our findings to date have led to the working hypothesis that SAFB1 down-regulation promotes a phenotype in CRPC that results in conservation of residual androgen in the tumor, thereby promoting an "intracrine" mechanism of AR activation. In contrast, SAFB1 appears to cooperate with EZH2 in silencing genes in a manner that results in a manner that opposes the AR activation signature that reflects the conventionally-understood pattern AR transcriptional activity. Interestingly, our data suggest that SAFB1 may cooperate with other proteins that act to *deplete* androgen from the tumor. This hypothesis is consistent with our bioinformatics analysis of thousands of RNA expression profiles from human prostate cancers that we have incorporated into this study. We find that about 1/3 of human prostate cancers, including primary tumors, exhibit an "AR activation suppressed" phenotype. We hypothesis that SAFB1 is a critical mediator of this phenotype. We will continue to test this novel idea in year 2 of this project.

Major accomplishments of this funding cycle include:

1) We identified the chromatin binding sites by SAFB1 by global analysis.

Identifying chromatin sites bound by SAFB1 in prostate cancer cells using chromatin immunoprecipitation and next generation DNA sequencing (ChIP-seq): Due to the limited binding affinity of SAFB1 antibody (Sigma-Aldrich), endogenous SAFB1 binding DNA fragments could not be enriched for ChIP-seq analysis. Thus, in order to increase precipitation efficacy on SAFB1, a SAFB1 expressing vector construct with an HA tag was transfected into LNCaP cells and precipitated with HA tag antibody for ChIP-seq library. To characterize the SAFB1 cistrome, chromatin sites bound by SAFB1-HA were identified using ChIP-seq. LNCaP cells were treated with 1 nM DHT or vehicle and chromatin immunoprecipitation was performed with HA tag antibody at 4 hour time points using a ChIP protocol. ChIP DNA was converted into libraries and sequenced using the Illumina HiSeq2000.

Conducting computational analysis of ChIP-seq data for SAFB1 cistrome: For identification of chromatin binding sites of SAFB1, sequencing data was processed using the Illumina analysis pipeline, aligned to the UCSC hg19/NCBI 37 version of the human genome using Bowtie[7], reads with the exact same mapping location were considered to be PCR duplicates and collapsed into a single record using samtools[8], and SAFB1-enriched binding sites were identified using the R csaw package[9]. ***17,884 genome-wide SAFB1 binding sites on promoter regions (upstream 2,000 bp and downstream 500 bp from transcription start sites (TSS)) were identified by comparing with binding sites in input control sample.*** Overlap and feature annotation of ChIP-seq enriched regions were performed using R detailRanges function from csaw package[9]. Intersecting with 922 differentially expressed genes (DEGs) by SAFB1 knockdown, I found that 259 DEGs contain SAFB1 binding sites in their promoter regions. This result suggests that about 28% of DEGs can be regulated by SAFB1 binding in their proximal promoters. Notably, steroid and androgen metabolism related genes ***(ASMTL, CYP21A2, UGT2B15, UGT2B17, and HSD17B8)*** were identified. This result is highly consistent with our preliminary data showing massive down-regulation of sterol metabolism genes with SAFB1 silencing.

2) We determined the effects of SAFB1 knockdown on the AR and EZH2 cistromes.

Performing ChIP-seq using anti-specific antibodies against AR or EZH2: ChIP-seq analysis was performed to identify AR and EZH2 target genes dependent on SAFB1 loss in LNCaP cells in the presence or absence of DHT. LNCaP SAFB1 knockdown and control cells were treated with 1 nM DHT and chromatin immunoprecipitation was performed with AR and EZH2 antibody at 4 hour time points. ChIP DNA was converted into libraries and sequenced using the Illumina HiSeq2000.

Conducting computational analysis of ChIP-seq data for AR and EZH2 cistrome: ChIP-seq reads were mapped to the UCSC hg19/NCBI 37 version of the human genome using Bowtie[7]. Differential AR binding sites between the SAFB1 knockdown LNCaP cells and the control cells were found by using R csaw package[9]. As an additional filter, low-abundance windows contain no binding sites were filtered out. This improves power by 1) removing irrelevant tests prior to the multiple testing correction; 2) avoiding problems with discreteness in downstream statistical methods; and 3) reducing computational work for further analyses[10]. Filtering is performed using the average abundance of each window. Binding sites are only retained if they have abundances 10-fold higher than the background. This removes a large number of binding sites that are weakly or not marked and are likely to be irrelevant.

In order to compare the list of genes associated with AR and EZH2 binding peaks to the list of genes differentially expressed on SAFB1 knockdown, the list of gene symbols for promoters associated with AR or EZH2 binding sites were intersected with the list of gene symbols for DEGs. As a result, the AR and EZH2 ChIP-seq data in SAFB1 knockdown LNCaP cells produced 7,193 and 8,038 sites compared to control cells, respectively. 922 DEGs in SAFB1 knockdown were intersected with those genes identified as having a proximal or nearby AR and/or EZH2 binding sites in the either knockdown cells or control cells. ***PSMB8, HLA-B, UGT2B10, UGT2B15, KLK3 (PSA), IRS1, ABCF1, FDPS, ONECUT2, and SAFB1 were identified with significant increased or decreased binding (>1.5 fold) of AR in their promoter regions (Figure 1).*** This data suggests that there is a close relationship between SAFB1, AR, and EZH2 binding, and gene expression; on SAFB1 knockdown, genes associated with an AR binding peak are significantly more likely to be differentially expressed than other genes.

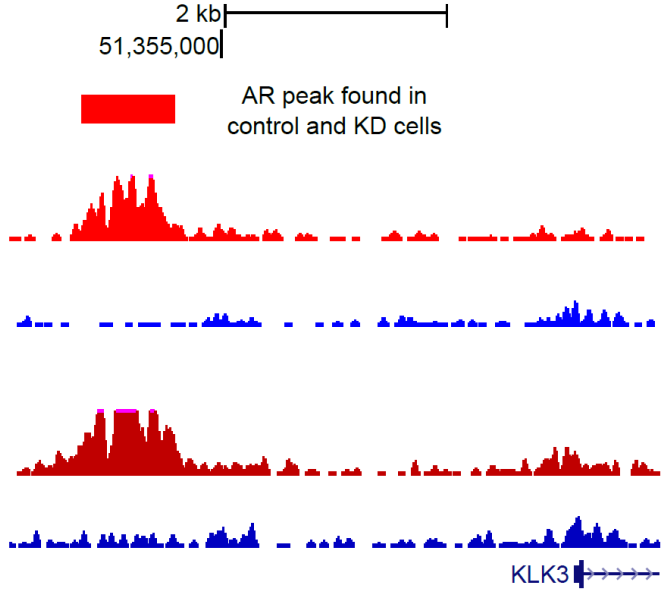


Figure 1. AR peaks on proximal promoter of KLK3 gene also known as prostate specific antigen (PSA) in both control and SAFB1 knockdown cells.

Validation of genes with AR and/or EZH2 binding sites identified in SAFB1 knockdown cells: LNCaP cells with the stable knockdown of SAFB1 using shRNA from Sigma Aldrich. The cells were analyzed for SAFB1 loss and amplified for AR protein expression and AR transcriptional activity[6].

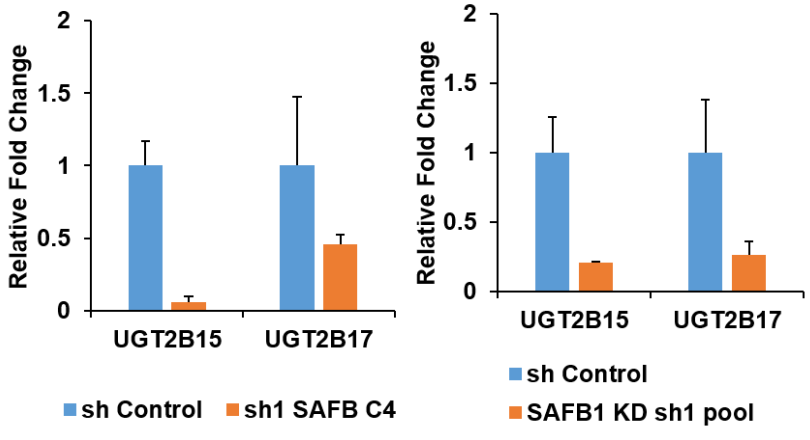


Figure 2. Decrease of UGT2B15 and 17 gene expression by SAFB KD in LNCaP (left) and 22Rv1 (right).

These SAFB1 knockdown cells had a downregulation of the UGT2B family of genes and UGT2B15 and UGT2B17 (Figure 2), the most well studied UGT2B genes within the prostate, were validated by qPCR using primers generated by Ohno et al.[11] and shown to be specific for the different family members. Applied Biosystems ABI Prism 7900HT qPCR machine was used to perform the analysis.

UGT2B15 and 17 gene expression changes by SAFB1 overexpression (OE) were analyzed in LNCaP cells. qPCR analysis of UGT2B15 and 17 in LNCaP cells was done after transient OE of SAFB1-HA tag (pBABE vector backbone) for 48 hours using Lipofectamine LTX (Invitrogen). Overexpression of SAFB1 was confirmed by qPCR (Figure 3A). Then, qPCR analysis was performed to measure UGT2B15 and 17 gene expression changes by SAFB1 OE. The qPCR primers and the qPCR equipment are the same as above. ***Significant increase of UGT2B15 and 17 gene expression were confirmed (Figure 3B).*** To validate whether this expression changes are directed by SAFB1 binding in promoter regions of UGT2B15 and 17 genes, we performed ***luciferase analysis of UGT2B15 and 17 activity in 22Rv1 prostate cancer cells (Figure 3C)***. For this analysis, the UGT2B15 and 17 luciferase promoter constructs (PGL4.10 vector backbone from promega) were co-transfected with control or SAFB1-HA overexpression vector into 22RV1. Baseline activity was generated from empty luciferase PGL4.10, data was normalized to this negative control (set at 100% activity). The result shows that significant increase of promoter activity of UGT2B15 and 17 genes by SAFB OE. ***This result demostrates that UGT2B15 and 17 genes are regulated by binding of SAFB1 in the promoter region.***

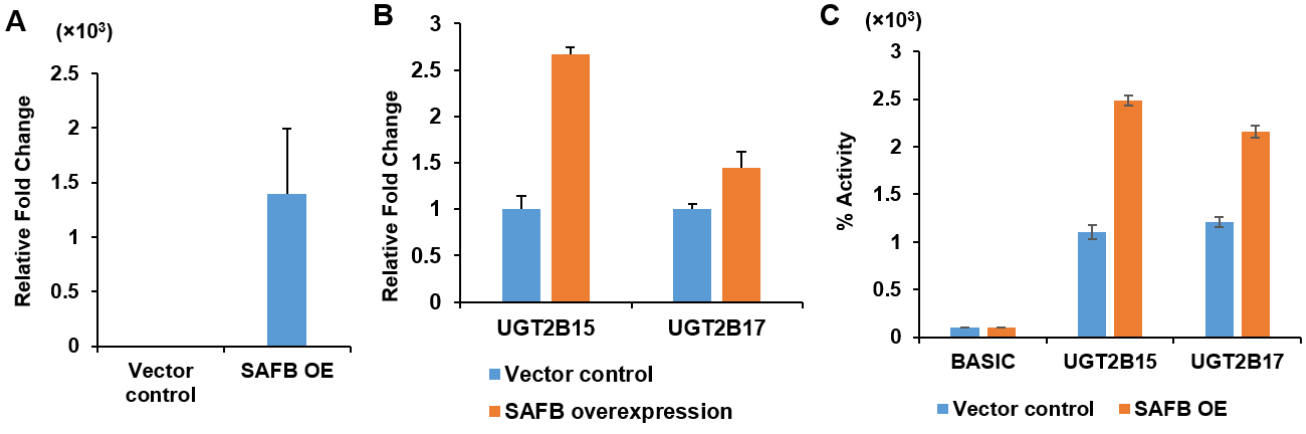


Figure 3. UGT2B15 and 17 gene expression changes and promoter activities perturbed by SAFB1 overexpression (OE) in LNCaP and 22Rv1. (A) Overexpression of SAFB1 in LNCaP. (B) qPCR analysis of UGT2B15 and 17 genes perturbed by SAFB1 overexpression in LNCaP. (C) Luciferase activity of UGT2B15 and 17 gene promoter by SAFB1 in 22Rv1.

<u>Identifying consensus binding motifs of SAFB1 and AR:</u> To identify a consensus binding motif for SAFB1 and AR, motif analysis was done for SAFB1 and AR ChIP-seq data sets. ***I found the AR/PR motif (Figure 4A) in ~60% of peaks and the AR half-site motif (Figure 4B) in ~75% of peaks.*** Figure 4 shows motif logos of SAFB1 and AR bining motifs from MEME analysis[12]. This result obtained from 1,069 common sitess between the SAFB1 and AR data sets.

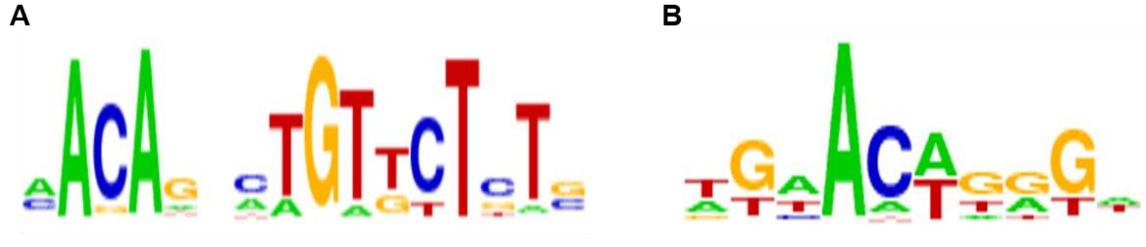
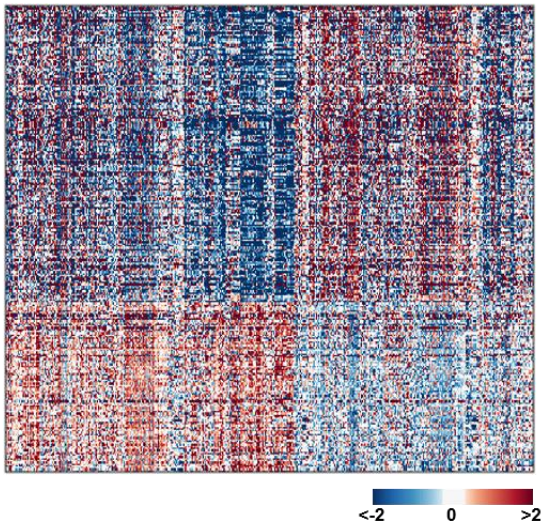


Figure 4. Consensus binding motifs of SAFB1 and AR. (A) AR/PR bidning motif. (B) AR half-site motif.

3) We found a clinical correlation of SAFB1 loss and PC progression and patient outcomes.

Transcriptome analysis revealed SAFB1 loss-dependent genes and pathways in clinical specimens: In order to identify genes in human PC tumors that correlate with alterations in SAFB1 gene expression, I compared 726 prostate tumor samples with low (<25 percentile) vs. high (>75 percentile) expression of SAFB1. Over 3,000 differentially expressed genes (DEGs) between prostate tumors with low (or no) and high expression of SAFB1 were selected with false dicovery rate (FDR)<0.05, and applied to functional enrichment analysis using DAVID software (Figure 5A and B). Enriched cellular processes indicate that SAFB1-dependent differential expression results in a more aggressive phenotype, including increases in steroid hormone reponses, regulation of blood vessel formation (angiogenesis), and regulation of RNA processing (Figure 5B). By integrating SAFB1 and AR ChIP-seq data sets and differential expression of SAFB1 knockdown cells with differential expression of prostate cancer patients, 387 genes were identified. These genes are differentially expressed in both SAFB1 knockdown cells and prostate cancer patients with low SAFB1 expression, as well as have SAFB1 and/or AR binding in their promoters. ***Among these genes, PSMB8, HLA-B, UGT2B10, UGT2B15, KLK3 (PSA), IRS1, ABCF1, FDPS, ONECUT2, and SAFB1 were also identified.***

A



B

Class	Term	P Value
Up-regulated in low SAFB	response to steroid hormone stimulus	3.42E-06
	regulation of tube size	1.38E-05
	regulation of blood vessel size	1.38E-05
	circulatory system process	1.69E-05
	blood circulation	1.69E-05
Down-regulated in low SAFB	RNA processing	1.09E-13
	mRNA processing	6.55E-13
	mRNA metabolic process	6.55E-13
	RNA splicing	1.09E-12
	RNA splicing, via transesterification reactions	2.16E-09

Figure 5. Differentially expressed gene by SAFB1 loss in clinical samples and their enriched cellular processes. (A) Heatmap depicts differential expression pattern of SAFB1 dependent gene signnature in prostate cancer. (B) Enriched cellular processes by up- or down-regulated genes between patients with SAFB1-high (>75 percentile) and SAFB1-low (<25 percentile).

Development of a new classification scheme for prostate cancer: To better understand the molecular heterogeneity of prostate cancer, I have assembled a "prostate cancer transcriptome atlas" (PCTA) software tool and database that contains more than 4,000 human prostate cancer transcriptomes assembled from public databases and the literature (including GEO, Array Express and TCGA). Using the PCTA, I examined transcriptome-based patterns of diverse oncogenic pathways and other important features in prostate cancer using a collection of 22 previously published gene expression signatures[13-29], resulting in a summary of activity score data of 14 pathways of the tumors. When applied an unsupervised clustering algorithm based on non-negative matrix factorization (NMF)[30] to pathway activity score data consisting of 1,321 prostate tumors, I identified three distinct sub-groups, shown below as Group 1-3 (Figure 6).

The heatmap in Figure 6 shows the surprising result that identifiable molecular features are evident across all disease categories through Gleason Score (GS) <7 to metastatic or castration-resistant prostate cancer (CRPC/Met), suggesting that prostate tumors retain identifiable epigenomic properties as tumor evolution proceeds. Although there are exceptions to the broad patterns, we found

a remarkable consistency within groups. ERG fusion-inducible gene expression is predominant in Group 1, which is also characterized by high AR activation activity scores. AR-variant inducible gene expression is clustered in Group 2, which also shows high proliferation and neuroendocrine activity. In contrast to the features seen in Group 1 and Group 2, Group 3 is uncharacterized as a distinct entity in prostate cancer. Group 3 exhibits pro-neural and mesenchymal activation signatures. Notably, the AR activation signature and AR variant-inducible signatures are relatively low in Group 3.

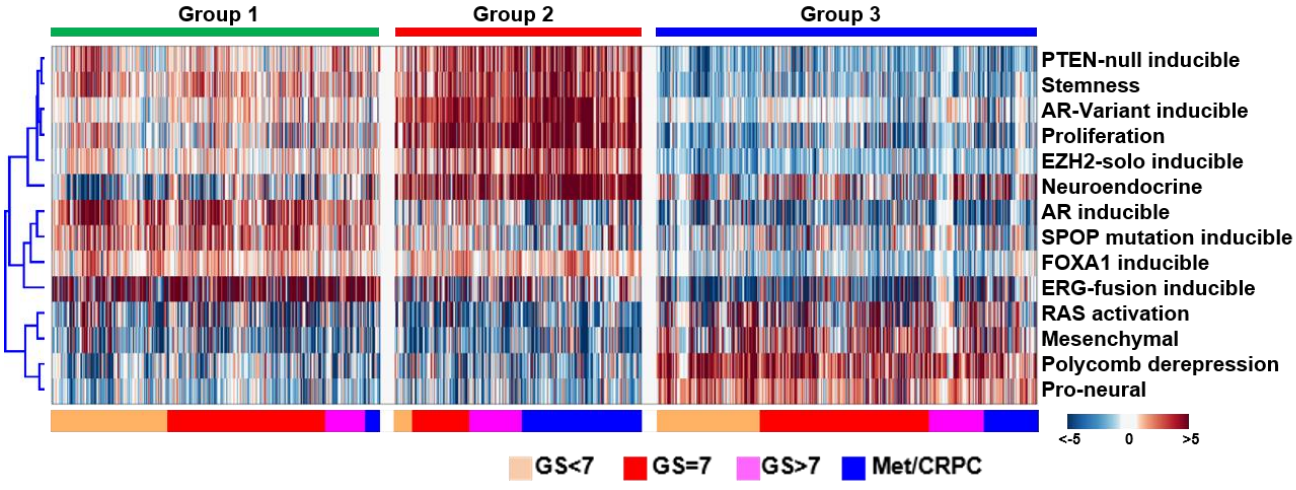


Figure 6. Patterns of signature pathway activities of 1321 prostate cancer patients in transcriptome atlas. (A) Patterns of activity scores were determined for each sample using Z score method. Consensus NMF clustering of 1321 prostate tumors using 14 pathway activity scores revealed three intrinsic molecular subtypes of prostate cancer (Group 1-3). The pathway activity scores (y-axis) were clustered by complete linkage hierarchical clustering method.

Validation of the three subtypes using independent cohorts: I have validated this classification system in 10 independent patient series, consisting of over 1,200 RNA expression profiles. This result ***suggests that it might be possible to cluster essentially all prostate cancers into one of only 3 subtypes*** defined by gene signatures that have been functionally implicated in the disease.

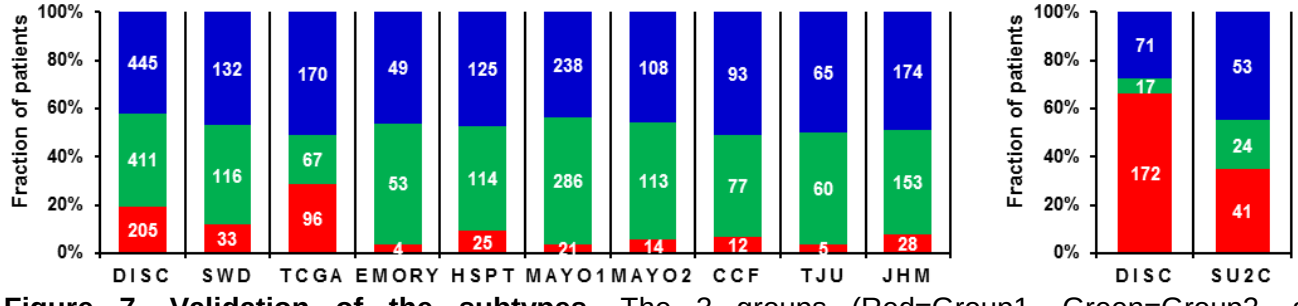


Figure 7. Validation of the subtypes. The 3 groups (Red=Group1, Green=Group2, and Blue=Group3) were recognized in 10 independent cohorts. Comparable fractions of patients with primary prostate tumors (left) and CRPC/Met (right) are assigned to each subtype within the different cohorts. DISC=discovery cohort; SWD=Swedish watchful waiting cohort; TCGA=TCGA cohort; EMORY=Emory cohort; HSPT= Health Study Prostate Tumor cohort; MAYO1=Mayo clinic cohort 1; MAYO2=Mayo clinic cohort 2; CCF=Cleveland clinic cohort; TJU=Thomas Jefferson University cohort; SU2C= SU2C/PCF Dream Team cohort.

Discovering a novel driver of aggressive prostate cancer variants: To computationally identify transcription factors (TFs) that are highly active in this disease space, I used the large number of CRPC/Met tumors (n=260) in the PCTA. We integrated RNA expression data with TF-target gene interaction data collected from a number of chromatin immunoprecipitation (ChIP) and curated databases that contain genes that share TF binding sites. We then conducted a master regulator analysis (MRA) based on a combination of gene set enrichment analysis (GSEA) and rank correlation of TF expression level and RNA expression level of known targets for each TF. This analysis identified

a set of TFs known to be functionally significant in CRPC/Met prostate cancer, including AR, EZH2, FOXM1, and E2F3, thereby validating our approach. Surprisingly, this analysis also identified a TF that has not been studied in prostate cancer, ONECUT2, an atypical homeobox TF that has been implicated in liver, pancreas and neural development. Notably, ONECUT2 is one of SAFB1 target genes in prostate cancer cells. We found that ONECUT2 gene expression is significantly down-regulated by SAFB1 KD in LNCaP (Figure 8). ONECUT2 expression gradually increases across the disease categories in the PCTA data set, and bionformatics modeling predicts that it functionally interacts with AR, EZH2, and FOXA1. Enforced and silenced ONECUT2 in LNCaP and 22Rv1 cells were done by transfecting shONECUT2 and ONECUT2-overexpressing vector construct, and conducted oligonucleotide expression array and functional experiments. Significantly, ONECUT2 can potently inhibit AR, PSA, EZH2, and FOXA1 expression (Figure 8), consistent with our computational modeling predictions.

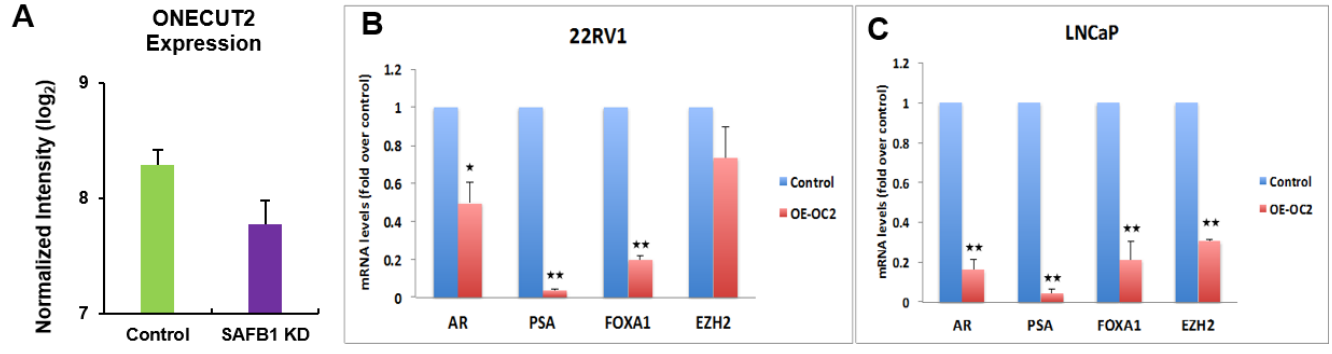


Figure 8. Gene expression of ONECUT2, AR, PSA, FOXA1 and EZH2. (A) Differential gene expression of ONECUT2 gene by SAFB1 knockdown in LNCaP (FDR<0.05). (B) ONECUT2 (OC2) suppresses AR, PSA, FOXA1, and EZH2 in prostate cancer cell lines. Total RNA was isolated from 22Rv1 and LNCaP cells overexpressing OC2 and real-time qPCR was performed using TaqMan probes for the indicated genes. Each value represents the mean±SEM of 3 independent experiments performed in triplicate. Significant differences are denoted by asterisks (*p≤0.05. **p≤0.01).

ONECUT2 plays a role in stimulating growth of 22Rv1 cells (Figure 9A) and therefore might be targeted in vivo to limit progression of CRPC. The PSA/KLK3 enhancer is a prostate regulatory element, strongly supporting the role of ONECUT2 in prostate cancer. From this result, I hypothesized that ***ONECUT2 is a driver of PSA-negative clones that may expand after therapy and ONECUT2 expression level may be an indicator of progression to metastasis (Figure 9B).***

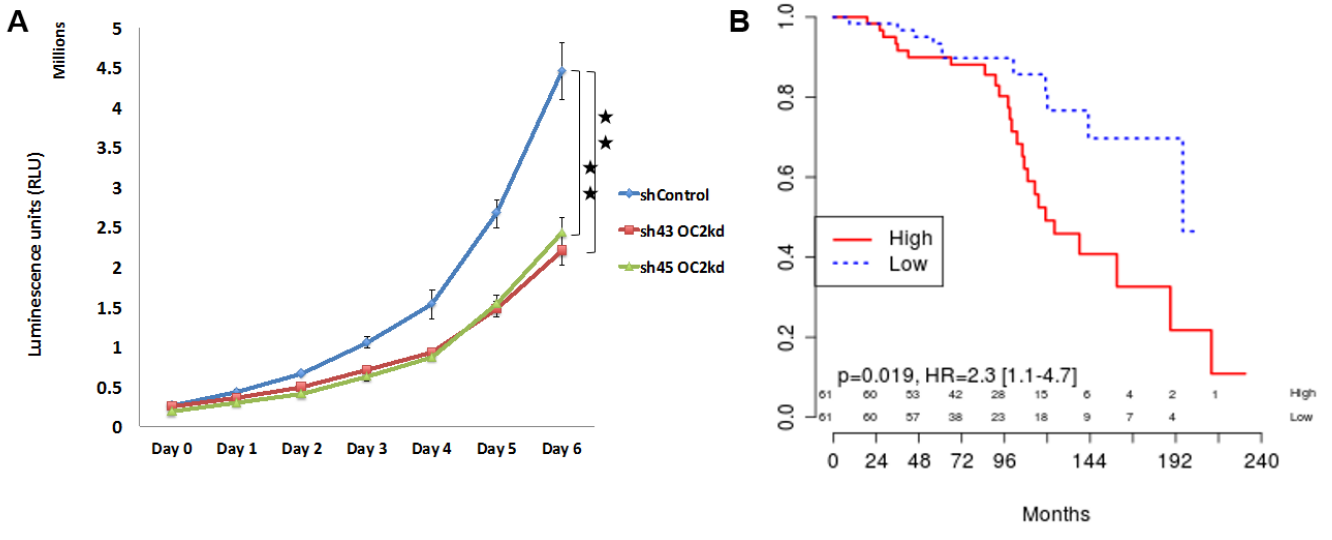


Figure 9. Biological and clinical implication of ONECUT2. (Left) Proliferation assay demonstrated significant inhibition of cell proliferation by ONECUT2 knockdown (kd). (Right) Kaplan-Meier analysis

showing top vs. bottom tertiles of OC2 expression level in relation to metastasis-free survival (CCF cohort).

I employed a prostate cancer classification scheme that I have developed in order to ask whether ONECUT2 activity segregates between Groups 1-3. We used the gene expression perturbation data generated using enforced and silencing methods to nominate ONECUT2 activation and repression signatures. We then applied these signatures to the three subtypes developed from the PCTA data. We found that the ONECUT2 activation signature is most active in Group 3 in all disease categories, but that the ONECUT2 repression signature increases progressively in Group 2, with highest activity in CRPC/Met tumors (Figure 10A and B). These findings demonstrate that we can map master regulator activity onto human prostate cancer by integrating our classification scheme with laboratory data.

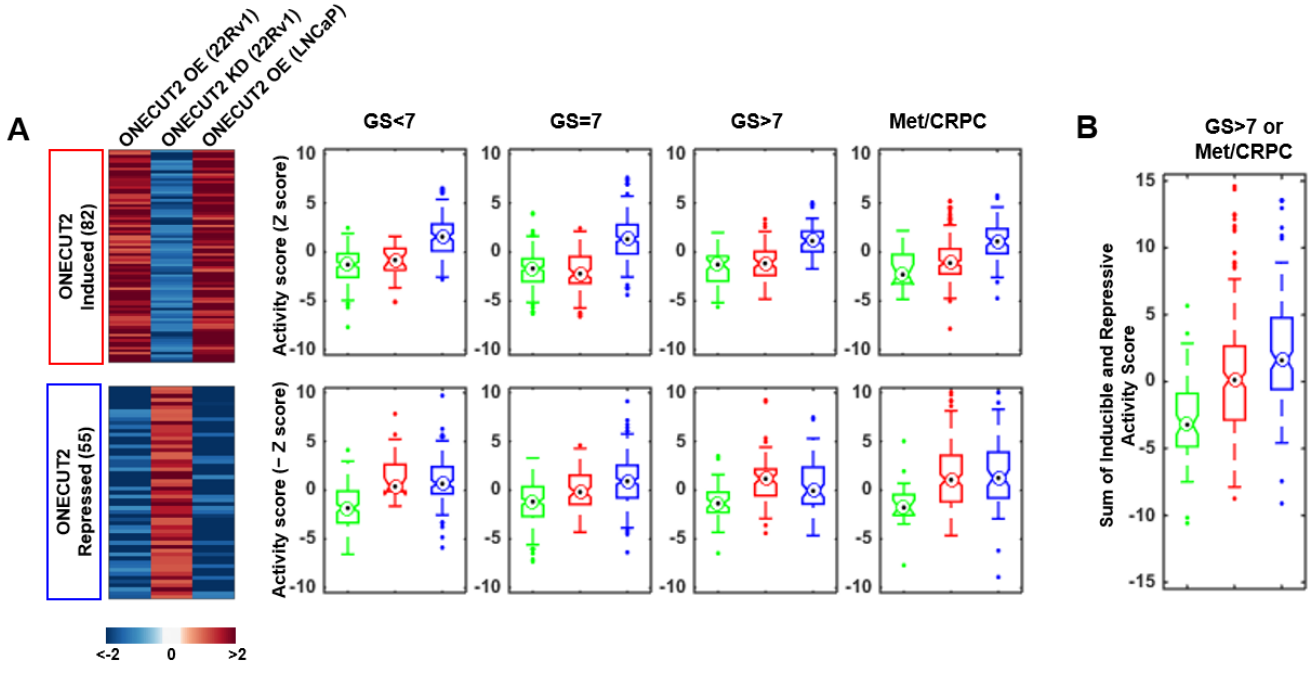


Figure 10. ONECUT2-inducible and -repressive activity is significantly enriched in group 3 in comparison to the other 2 groups. (A) Heatmaps show differential expression patterns of genes perturbed by ONECUT2 overexpression and knockdown in 22Rv1 and LNCaP cells (FDR<0.05 and fold change ≥2). Results from the PCTA cohort are shown in the panels A and B. Group 1 = green, group 2 = red, group 3 = blue.

4) Key research accomplishments

- Generation of the first ChIP-seq analysis of SAFB1 and the first identification of the SAFB1 cistrome in prostate cancer cells.
- Discovery that UGT2B15 and UGT2B17 are regulated by SAFB1, indicating that these androgen-inactivating genes are a component of the SAFB1 transcriptional network.
- Development of a novel classification system for prostate cancer that has utility in providing novel and actionable clinical information.
- Identification of ONECUT2 as a novel driver of aggressive prostate cancer variants.

5) Conclusion

ChIP-seq analysis followed by computational analysis permitted the determination of the extent to which chromatin occupancy of SAFB1 cistrome components reflects gene expression patterns characteristic of AR and EZH2 activity. Integration of our own ChIP-seq data and patient gene expression profiles allowed us to identify the extent of transcriptional collaboration of SAFB1, AR, and EZH2 in prostate cancer cells and human prostate tumors. UGT2B15/17 genes demonstrated to reside within this AR-metabolic network will represent a high priority for further study.

6) Other achievement

We found that UGT2B15/17 gene expression is significantly increased in androgen independent LNCaP clones (C-81). Using this system, we have developed a platform to measure metabolic changes (such as DHT) by modulation of metabolic genes regulated by SAFB1

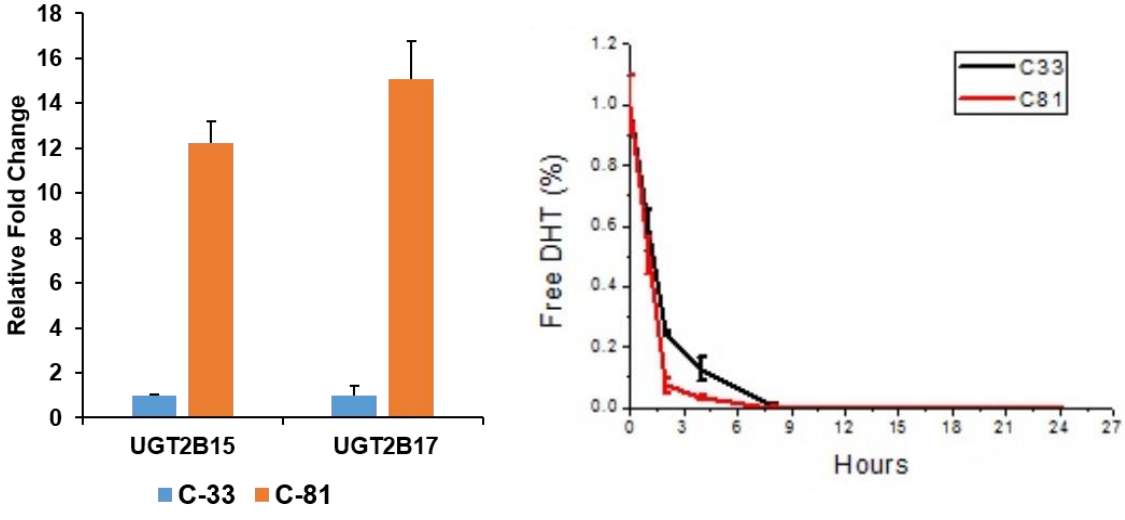


Figure 11. Development of measure of metabolite changes. (Left) LNCaP derived C33 and C81 cells generated by the lab of Min Fong Lin are a cell line that were passaged in 5% fetal bovine serum for 33 passage (C33) and 81 passages (C81). These C81 cell line are hormone insensitive while C33 are still hormone sensitive. Upon comparing these cells, there is a large increase ~10-15 fold) in UGT2B15 & 17 in the hormone insensitive cell line C81 in comparison to hormone sensitive C33. (n=3). (Right) LNCaP derived C33 and C81 cells were analyzed by HPLC in the lab of Nima Shariffi. C81 has a (slightly) stronger activity as free DHT signal decreased faster at 2 hours. The method employed was to treat 1 million cells with 100nM cold plus some hot DHT, and examine hydrophobic radioactive signals (majorly DHT) in culture media by HPLC. (n=3)

7) Reference

1. Solomon KR, Freeman MR. Do the cholesterol-lowering properties of statins affect cancer risk? Trends Endocrinol Metab. 2008;19: 113-121.
2. Platz EA, Leitzmann MF, Visvanathan K, et al. Statin drugs and risk of advanced prostate cancer. J Natl Cancer Inst. 2006;98: 1819-1825.
3. Flick ED, Habel LA, Chan KA, et al. Statin use and risk of prostate cancer in the California Men's Health Study cohort. Cancer Epidemiol Biomarkers Prev. 2007;16: 2218-2225.
4. Murtola TJ, Visakorpi T, Lahtela J, Syvala H, Tammela T. Statins and prostate cancer prevention: where we are now, and future directions. Nat Clin Pract Urol. 2008;5: 376-387.
5. Hamilton RJ, Banez LL, Aronson WJ, et al. Statin medication use and the risk of biochemical recurrence after radical prostatectomy: results from the Shared Equal Access Regional Cancer Hospital (SEARCH) Database. Cancer. 2010;116: 3389-3398.
6. Mukhopadhyay NK, Kim J, You S, et al. Scaffold attachment factor B1 regulates the androgen receptor in concert with the growth inhibitory kinase MST1 and the methyltransferase EZH2. Oncogene. 2013.
7. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10: R25.
8. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078-2079.
9. Lun AT, Smyth GK. De novo detection of differentially bound regions for ChIP-seq data using peaks and windows: controlling error rates correctly. Nucleic Acids Res. 2014;42: e95.

10. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40: 4288-4297.
11. Ohno S, Nakajin S. Determination of mRNA expression of human UDP-glucuronosyltransferases and application for localization in various human tissues by real-time reverse transcriptase-polymerase chain reaction. Drug Metab Dispos. 2009;37: 32-40.
12. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43: W39-49.
13. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell. 2006;9: 157-173.
14. Beltran H, Rickman DS, Park K, et al. Molecular characterization of neuroendocrine prostate cancer and identification of new drug targets. Cancer Discov. 2011;1: 487-495.
15. Assou S, Le Carrour T, Tondeur S, et al. A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells. 2007;25: 961-973.
16. Yu J, Yu J, Rhodes DR, et al. A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. Cancer Res. 2007;67: 10657-10663.
17. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439: 353-357.
18. Lee TI, Jenner RG, Boyer LA, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell. 2006;125: 301-313.
19. Xu K, Wu ZJ, Groner AC, et al. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science. 2012;338: 1465-1469.
20. Geng C, Rajapakshe K, Shah SS, et al. Androgen receptor is the key transcriptional mediator of the tumor suppressor SPOP in prostate cancer. Cancer Res. 2014;74: 5631-5643.
21. Robinson JL, Hickey TE, Warren AY, et al. Elevated levels of FOXA1 facilitate androgen receptor chromatin binding resulting in a CRPC-like phenotype. Oncogene. 2014;33: 5666-5674.
22. Jin HJ, Zhao JC, Ogden I, Bergan RC, Yu J. Androgen receptor-independent function of FoxA1 in prostate cancer metastasis. Cancer Res. 2013;73: 3725-3736.
23. Iljin K, Wolf M, Edgren H, et al. TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. Cancer Res. 2006;66: 10242-10246.
24. Setlur SR, Mertz KD, Hoshida Y, et al. Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. J Natl Cancer Inst. 2008;100: 815-825.
25. Nelson PS, Clegg N, Arnold H, et al. The program of androgen-responsive genes in neoplastic prostate epithelium. Proc Natl Acad Sci U S A. 2002;99: 11890-11895.
26. Sharma NL, Massie CE, Ramos-Montoya A, et al. The androgen receptor induces a distinct transcriptional program in castration-resistant prostate cancer in man. Cancer Cell. 2013;23: 35-47.
27. Mendiratta P, Mostaghel E, Guinney J, et al. Genomic strategy for targeting therapy in castration-resistant prostate cancer. J Clin Oncol. 2009;27: 2022-2029.
28. Tomlins SA, Laxman B, Varambally S, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia. 2008;10: 177-188.
29. Hu R, Lu C, Mostaghel EA, et al. Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer. Cancer Res. 2012;72: 3457-3462.
30. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A. 2004;101: 4164-4169.

What opportunities for training and professional development has the project provided?

I have been promoted to the rank of Instructor at Cedars-Sinai, a position from which I can submit independent grant proposals. I have also been a co-author of 6 published studies (see PRODUCTS section). I also gave a poster presentation at the 2014 Annual Conference of the American Urological Association (AUA) and the 2014 Prostate Cancer Foundation (PCF) meeting (as shown in PRODUCT). To support education and teaching of bioinformatics and computational methods within the Cedars-Sinai prostate cancer research community, I gave presentations in lab meetings, journal club, and cancer biology workgroup meetings. I have substantive one-to-one discussions with the mentors several times per week and is in near-constant contact via e-mail, telephone or Skype. I

have (and will continue) close communication with other senior investigators through many other routes, including (1) weekly joint lab meetings, (2) bi-weekly Cancer Biology Journal Club (organized by Dr. Kim), and (3) bi-weekly Cancer Genomics Journal Club (organized by Dr. Kim). This is a very interactive community with open lines of communication across 8 nationally prominent prostate cancer research laboratories, where opinions, regents and data are continuously shared.

At this stage in the funding cycle, I have submitted a paper that is under review describing the new prostate cancer classification scheme that we have developed and its possible clinical significance, which was DIRECTLY derived from this proposed study.

How were the results disseminated to communities of interest?

I created a large (>4,000 specimens) RNA expression data set from prostate cancer and benign prostate tissue. From this large data set, he has demonstrated for the first time that prostate cancers, possibly all prostate cancers, can be subtyped to only three distinct groups. This is a major discovery in the field, which has allowed collaborations with other nationally prominent prostate cancer research teams, including at the University of Michigan and UCLA. Early versions of this work have also been presented at AUA and PCF national conferences. These findings are being prepared for publication now.

What do you plan to do during the next reporting period to accomplish the goals?

A major objective of the second year of funding period is to test whether cholesterol alters intracrine androgen mechanisms in a SAFB1-dependent manner. I will measure expression of AR-regulated genes relevant to PC using qRT-PCR and will perform Metabolite profiling using mass spectrometry coupled to bioinformatics analysis. In addition, I will characterize the function of ONECUT2 in prostate cancer progression. To this end, I will apply a set of experimental and bioinformatics strategies to understand the function of ONECUT2 in prostate cancer, and develop approaches directed toward targeting it.

4. IMPACT

What was the impact on the development of the principal discipline(s) of the project?

I have made an important conceptual and clinically relevant advance by developing a novel method of characterizing prostate cancer using transcriptomic profiles. Consequently, this project is high impact and high reward, with potentially immediate opportunities to alter clinical practice if the classification scheme can be shown to have clinical utility. This new prostate cancer classification scheme I developed might improve prognostication of prostate cancer and enable the development of subtype-specific therapies and companion diagnostics. Using computational modeling, I have also identified a transcription factor, ONECUT2, which appears to be highly active in CRPC/Met tumors, but which has not been studied in prostate cancer, and therefore represents a first-in-field discovery.

What was the impact on other disciplines?

Nothing to Report.

What was the impact on technology transfer?

Nothing to Report.

What was the impact on society beyond science and technology?

Nothing to Report.

5. CHANGES/PROBLEMS

Changes in approach and reasons for change

Nothing to Report.

Actual or anticipated problems or delays and actions or plans to resolve them

There was a delay for submitting a paper that describes a result of this project to a high impact journal as shown in Journal Publications Section. I also have performed additional experiments and analyses to achieve the tasks on original SOW and put the results into this revised annual report. Due to this delayed report, revised SOW with adjusted target months to achieve major goals of the second year of funding period is provided in the Appendices Section.

Changes that had a significant impact on expenditures
Nothing to Report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
Nothing to Report.

6. PRODUCTS:

Publications, conference papers, and presentations

Journal publications.
You S, Knudsen BS, Erho N, Alshalalfa M, Takhar M, Ashab HA, Davicioni E, Karnes RJ, Klein EA, Den RB, Ross AE, Schaeffer EM, Garraway IP, Kim J, Freeman MR. Three intrinsic subtypes of prostate cancer with distinct pathway activation profiles differ in prognosis and treatment response. **Submitted.** Acknowledgement of federal support (Yes)

Books or other non-periodical, one-time publications.
Nothing to Report.

Other publications, conference papers, and presentations.
Poster presentation:
1. **You S**, Kim J, Freeman MR, An epigenomic pathway from cholesterol to intracrine androgen. The 2014 American Urological Association (AUA) Annual Meeting, held in Orlando, Florida, from May 16 to 21, 2014.
2. **You S**, Kim J, Freeman MR, Prostate Cancer Classification Using a Transcriptome Atlas. The Prostate Cancer Foundation (PCF) 21st Annual Scientific Retreat, held in Carlsbad, California, October from 23 to 25, 2014.

Lecture:
You S, Kim J, Introduction to Bioinformatics. The Urologic Oncology Program, held in Cedars-Sinai Medical Center, Los Angeles, California, March 10, 2015.

Website(s) or other Internet site(s)
Nothing to Report.

Technologies or techniques
Nothing to Report.

Inventions, patent applications, and/or licenses
Nothing to Report.

Other Products
Nothing to Report.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	Sungyong You
Project Role:	Principal Investigator

Researcher Identifier:	yousung1
Nearest person month worked:	5
Contribution to Project:	Dr. You has performed all the works in computational analysis and experiments
Funding Support:	The Urology Care Foundation Research Scholar Program

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?
Nothing to Report.

What other organizations were involved as partners?
Nothing to Report.

8. SPECIAL REPORTING REQUIREMENTS: Nothing to Report.

9. APPENDICES: Revised SOW

<u>Statement of Work</u> – 20/11/2013
Proposed Start Date 1 Sep, 2014

Site 1: Cedars-Sinai Medical Center [CSMC]
8700 Beverly Blvd.
Los Angeles, CA, 90048
PI: Sungyong You, PhD
Mentors: Michael R. Freeman, PhD
Jayoung Kim, PhD

Training-Specific Tasks:

Major Task 1: Training and educational development in prostate cancer research	**Months**	**CSMC**
Subtask 1: Fine-tune bench skills for experimental data generation and computational analysis methodology	1-12	Dr. You
Subtask 2: Attend weekly group meetings with other lab members and weekly one-on-one project meetings with two mentors	1-24	Dr. You
Subtask 3: Present research at Cancer Biology Program group meetings	1-24	Dr. You
Subtask 4: Present research at the monthly computational biology workgroup project meetings	1-24	Dr. You
Subtask 5: Attend bioinformatics, genomics, and cancer science minicourses and workshops.	1-24	Dr. You
Subtask 6: Attend a national scientific meeting in relevant	13-24	Dr. You

scientific field (e.g. AACR, IMPaCT)		
Milestone(s) Achieved: *Presentation of project data at a national meeting*	24	

Research-Specific Tasks:

Specific Aim 1: To characterize the SAFB1 cistrome in prostate cancer cells and to determine the metabolic and biologic effects of SAFB1 loss.		
Major Task 1: Global analysis of chromatin binding sites by SAFB1		
Subtask 1: Identify chromatin sites bound by SAFB1 in prostate cancer cells using ChIP-seq Cell lines used: LNCaP	1-9	Dr. You
Subtask 2: Conduct computational analysis of ChIP-seq data for SAFB1 cistrome Software used: BWA, MACS, R Cluster package, Galaxy, CEAS, GREAT, GSEA	10-12	Dr. You
Milestone(s) Achieved: *Characterization of SAFB1 cistrome in the presence- or absence of dihydrotestosterone (DHT)*	12	Dr. You
Major Task 2: Investigate effect of SAFB1 knockdown on the AR and EZH2 cistromes		
Subtask 1: Perform ChIP-seq using anti-specific antibodies against AR or EZH2 Cell lines used: SAFB1 KD cells or control LNCaP cells	1-6	Dr. You
Subtask 2: Conduct computational analysis of ChIP-seq data for AR and EZH2 cistrome Software used: BWA, MACS, R Cluster package, Galaxy, CEAS, GREAT, GSEA	6-15	Dr. You
Subtask 3: Perform qRT-PCR and ChIP-PCR to validate genes and chromatin binding sites identified in Subtask 2 Cell lines used: SAFB1 KD cells or control LNCaP cells	6-15	Dr. You
Subtask 4: Identify consensus binding regions of SAFB1, AR, and EZH2 Software used: MEME, DREME	6-15	Dr. You
Milestone(s) Achieved: *Determination of (1) the overlapping target genes or sub-network between SAFB1 and AR or EZH2; and (2) the genes or pathways involved in sterol metabolism and chromatin regulation.*	15	Dr. You

Major Task 3: Assess clinical correlation of SAFB1 loss and PC progression and patient outcomes		
Subtask 1: Assess the extent of correlation of SAFB1 loss with *patient outcomes* using bioinformatic approaches coupled to the Prostate Cancer Transcriptome Atlas (PCTA)	9-15	Dr. You
Subtask 2: Assess the extent of correlation of SAFB1 loss with *PC progression* using bioinformatic approaches coupled to the Prostate Cancer Transcriptome Atlas (PCTA)	9-15	Dr. You
Milestone(s) Achieved: *Determination of (1) the genes or pathways strongly associated with SAFB1 regulation and PC progression*	15	Dr. You
Specific Aim 2: To test whether cholesterol alters intracrine androgen mechanisms in a SAFB1-dependent manner		
Major Task 1: Test whether SAFB1 knockdown results in activation of an intracrine AR network arising from increased levels of intracellular androgen		
Subtask 1: Measure expression of AR-regulated genes relevant to PC using qRT-PCR Cell lines used: control and SAFB1 knockdown C4-2 cells	15-21	Dr. You
Subtask 2: Metabolite profiling using mass spectrometry coupled to bioinformatics analysis Cell lines used: control and SAFB1 knockdown C4-2 cells Software used: Pinpoint and Sieve software [Thermo Fisher]	15-21	Dr. You
Milestone(s) Achieved: *Identification of critical regulatory nodes in the androgen metabolism network*	21	
Major Task 2: Test whether an intracrine AR network responds to decreases and increases in cholesterol levels		
Subtask 1: Measure the levels of androgen-regulated genes and metabolites at 0, 4 and 16 hours after the switch into cholesterol-depleted medium (CDM) using qRT-PCR and MS/MS analysis Cell lines used: control and SAFB1 knockdown C4-2 cells Software used: Pinpoint and Sieve software [Thermo Fisher]	21-24	Dr. You
Subtask 2: Measure expression levels of androgen responsive genes, metabolites of androgen metabolism, cell proliferation and invasion in high cholesterol condition using qRT-PCR and MS/MS analysis Cell lines used: control and SAFB1 knockdown C4-2 cells Software used: Pinpoint and Sieve software [Thermo Fisher]	21-24	Dr. You

Milestone(s) Achieved: *(1) Characterization of the involvement of SAFB1 regulation of the the UGT2B gene family, androgen metabolism, and downstream effects relevant to disease progression; (2) publication of 1-2 peer reviewed papers*	24	